Final Technical Report

Project Title:

Exporting Real-Time GNSS Data and Products to the NEIC: Collaborative Research with University of Washington, Central Washington University, and US Geological Survey

USGS NEHRP Award numbers:

G19AP00017 (UW), G19AP00018 (CWU)

Principal Investigators:

Brendan W. Crowell
Department of Earth and Space Sciences
University of Washington
Johnson Hall Rm-070, Box 351310
4000 15th Avenue NE
Seattle, WA 98195-1310
(858) 750-8513
crowellb@uw.edu

Timothy I. Melbourne
Department of Geosciences
Central Washington University
400 E. University Way
Ellensburg, WA 98926-7417
(509) 963-2799
tim@geology.cwu.edu

Start and End Dates:

January 1, 2019 to December 31, 2019

This material is based upon work supported by the U.S. Geological Survey under Grant Nos. G19AP00017 and G19AP00018.

Abstract

In the minutes following large earthquakes, the inclusion of geodetic information is paramount to quickly characterizing the seismic source using information from the near-field. The goal of this NEHRP project is to provide the National Earthquake Information Center (NEIC) with an easy to use method of accessing and manipulating high-rate Global Navigation Satellite System data. This project consisted of two parts, first, Central Washington University processed data for over 800 stations globally and broadcast the position streams to the University of Washington over a RabbitMQ exchange. Secondly, the University of Washington built a dedicated server that placed all this information into a database that can be quickly queried using a set of Python scripts that the NEIC has full access to and can further build on top of. This database has been running almost continuously since August, 2019 with high data completeness (> 95%) and low latency (< 2 s). The system successfully recorded motions for several stations during the January 28, 2020 Mw 7.7 Lucea, Jamaica earthquake.

Introduction

The United States Geological Survey's (USGS) National Earthquake Information Center (NEIC) provides rapid estimates of earthquake size and impacts, publishes products such as ShakeMaps and PAGER alerts, and serves as one of the most reliable providers of earthquake information for significant events around the world. Information is published in minutes to tens of minutes, with updates occurring over days for significant events. The NEIC operates with a series of automated source products that are reviewed and refined by analysts over tens of minutes [e.g., *Hayes et al.*, 2011]. The first magnitude estimate is provided using the Mwp algorithm, followed by a series of W-phase CMT inversions [*Kanamori and Rivera*, 2008; *Duputel et al.*, 2011], a different CMT inversion using the full waveforms [*Polet and Thio*, 2011], and finally a finite fault inversion. Intertwined between these different models are both internal and external releases of information, ShakeMap and PAGER alert publishing, and coordination between other agencies published results (i.e. Pacific Tsunami Warning Center (PTWC)).

Unfortunately, near-field seismic data can suffer from a condition known as magnitude saturation, that is, as an earthquake gets larger and larger, it is difficult to statistically discern the magnitude from the P-wave information recorded on strong-motion and broadband instruments. A magnitude 7, 8, or 9 earthquake may be indistinguishable from one another and moreover, broadband instruments will clip in the near-field. Sensor rotations and tilts also make it difficult to accurately integrate acceleration into displacement without high-pass filtering, thus reducing the fidelity of the observations [*Melgar et al.*, 2013]. One of the most often cited magnitude saturation cases is the earthquake, and subsequent, tsunami warnings following the 2011 Mw 9.0 Tohoku-oki earthquake [*Hoshiba et al.*, 2011]. While the earthquake early warning system

worked well operationally by providing alerts prior to strong shaking, the magnitude estimate only reached M8.0 by 1 minute where it stayed for the duration of the early warning. This had a fairly pronounced impact on the ensuing tsunami warning where the maximum wave heights along the Sendai coast were underestimated by many meters, breaching the protective seawalls, and leading to many casualties. While it is difficult to estimate if there would be a difference in loss of life with a more accurate magnitude estimate due to evacuation protocols (policy might not dictate a difference in evacuation), obtaining a more accurate estimate of earthquake magnitude quickly could lead to a more nuanced emergency response. *Hayes et al.* [2011] outlined the NEIC's performance during the Tohoku-oki earthquake. The PTWC published an observatory message with Mwp = 7.5 at 4.7 minutes after origin time (OT). The first public release, which was a coordinated message between the JMA and PTWC was a Mw 7.9 at 18.6 minutes after OT. It should be noted the NEIC had an internal solution of Mw 8.5 at 8.3 minutes after OT, although this estimate slowly decreased in size by the time of the joint message. Internally, the W-phase vertical component solution was available at 19.6 minutes after OT, with the correct magnitude and focal planes roughly in line with the correct ones. However, the first ShakeMap and PAGER solution were released at 20.6 and 23.6 minutes respectively, based upon the smaller magnitude of 7.9. At 32 minutes, the correct W-phase result was released publicly and the other products were revised accordingly over the next 20 minutes. While these results are respectable given that all results are reviewed by analysts, reducing the impacts of magnitude saturation in the initial minutes will help to accelerate the timelines above.

The primary objective of our project is to provide the NEIC with an alternative magnitude and source estimate that does not saturate for large magnitude events. Global Navigation Satellite System (GNSS) methods do not suffer the problem of magnitude saturation due to the direct measurement of ground displacement [*Melgar et al.*, 2013]. While GNSS data processed in real-time has significantly greater noise than seismic instruments (1 cm horizontal, 3-5 cm vertical; *Genrich and Bock* [2006]), it has been shown to reliably capture earthquake ground motions above Mw 6 in the near-field [*Geng et al.*, 2013; *Melgar et al.*, 2015a]. It is because of this that GNSS data has been proposed to augment earthquake and tsunami early warning systems [*Blewitt et al.*, 2006; *Crowell et al.*, 2009]. Utilizing GNSS-derived peak ground displacements (PGD) for rapid magnitude estimation was first proposed by *Crowell et al.* [2013] and expanded by *Melgar et al.* [2015b] with a larger subset of earthquakes. The methodology is simple but powerful; using a log-log relationship between PGD and distance to the source (hypocentral or centroidal distance), magnitude can be robustly estimated for earthquakes ranging from M6-9. Furthermore, *Melgar et al.* [2015b] found that this estimate is attainable before the earthquake has ended, roughly halfway to two-thirds of the way through the source time function depending on the locations of the earthquake and stations. Since then, many other earthquakes have been validated using this approach such as the 2015 Mw 8.3 Illapel [*Crowell et al.*, 2018a], the 2016 Mw 7.1 Iniskin [*Grapenthin et al.*, 2017], the 2016 Mw 7.8

Kaikoura [*Crowell et al.*, 2018b] and the 2017 Mw 8.2 Tehuantepec (https://www.unavco.org/highlights/2017/chiapas.html).

The rapid estimation of coseismic displacements is useful for obtaining models of earthquake slip and allows for the derivation of ground motion and hazards impacts as well as tsunamigenesis [*Blewitt et al.*, 2006; *Crowell et al.*, 2009]. Obtaining even simplistic finite fault solutions will have significant impacts on ground motion prediction patterns, and thus, on derived ShakeMaps. Moreover, obtaining a reasonable slip model leads to a better model of seafloor deformation, and therefore, a more detailed view of tsunami inundation. In addition to static displacements, performing full dynamic inversions or joint inversions with GNSS displacements is advantageous to obtaining the full slip history, and therefore, a better model of strong ground motions. Providing the NEIC the full time series of deformation at near-field sites will allow them to include this information in their finite fault models in the future.

Data Processing at Central Washington University

CWU produces real-time position estimates at 1 Hz computed in the ITRF08 reference frame for stations across the West Coast of the United States using CWU's Fastlane positioning system [*Santillan et al.*, 2013]. Fastlane is currently a GPS-only system that estimates position using only carrier phase observables which greatly mitigates the influence of pseudorange multipath on position accuracy. Phase observables are internally continuously calibrated using geometry-free combinations of the L1 and L2 pseudorange in a Kalman-filter that simultaneously estimates the best floating point ambiguities while monitoring and correcting for possible cycle slips. Fastlane uses GPS carrier phase based only, unlike other precise point positioning (PPP) algorithms (e.g., *Kouba and Héroux* [2001]) that rely on both phase and pseudorange, and relies on the fact that the calibration procedure greatly mitigates the influence of pseudorange multipath that may affect the estimation of the floating point ambiguities. Using only half the number of input observations reduces the overall computation burden of position estimation and translates into smaller latencies. Fastlane employs carrier phase-only-based positioning rather than pseudorange because phase alone is far less contaminated by multipath error, one of the largest sources of noise in HR positioning. Fastlane depends on a highly efficient algorithm for the resolution of carrier phase initial ambiguities, which for most stations can be initially resolved in 20– 30 s. The resulting positions show a typical RMS scatter of 2.5 cm in the horizontal and 5 cm in the vertical with latencies below 2 seconds.

Database at University of Washington

The position estimates from CWU are multi-casted over a RabbitMQ exchange in geoJSON format. The information provided in the message is the 4-character station name, the observation time (in Unix time), the International Terrestrial Reference Frame (ITRF) Cartesian

coordinates of the station, the 6 components of the covariance matrix, and some additional metadata. A server operated at UW has two system services listening to this RabbitMQ exchange, both written in Python3. The first service appends the position, timing, and covariance information into text files which are available over the web. This service also creates a separate file to record the total latency of the observation as it arrives at UW. The latency in general is less than 2 seconds, which includes transmission time from station to CWU, processing time at CWU, and rebroadcast time to UW. At present, we have archived over 130 days of data as of early February 2020 for over 800 stations (the first archived date is September 24, 2019). The UW server has a 2-TB RAID drive as well as two 1-TB backup archives to store these text files. The second system service listening to the RabbitMQ exchange places the position solutions into a MongoDB database. This database creates a new collection for any new station that is received, and all positions are stored in these collections. We keep a rolling buffer of 5 days, after which, accessing data requires checking the text files.

Two crontab jobs are run every hour on the database. First, a metadata script runs over the entire database, determining the amount of available data there is for the past 24 hours, and records the station metadata into a JSON file. This JSON file is then used to populate a Leaflet map that displays the available stations, color codes them by data availability, and also plots the USGS significant earthquakes during the past week. Figure 1 shows a screenshot of this Leaflet map. The second crontab hourly job simply deletes any data that is older than 5 days. This is done to reduce the memory burden on the system; MongoDB stores all information in local memory to make queries significantly faster. With 5 days of information for 800 sites, we allocate roughly 30 GB of RAM to MongoDB out of 128 GB available on our server.

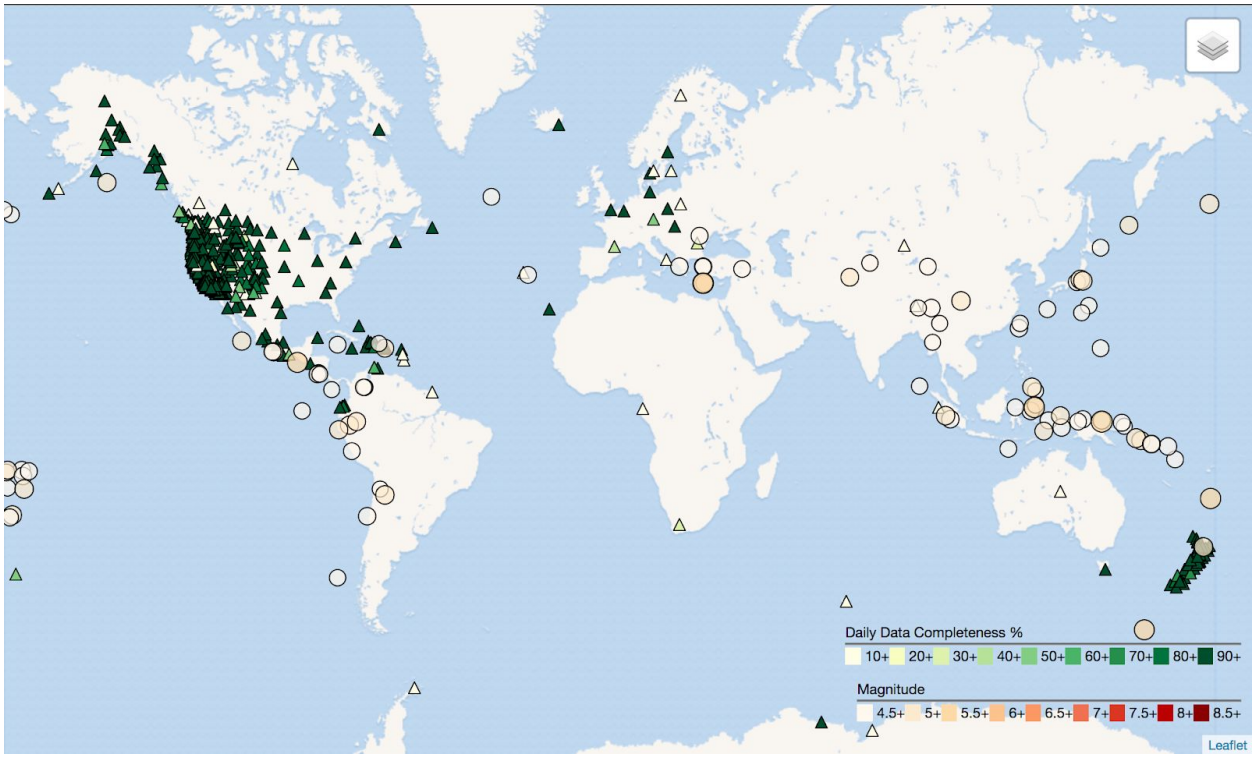

Figure 1: Current stations available as of February 5, 2020 in database at University of Washington. The triangles indicate stations that have data for the last 5 days and the coloration shows the last day's data completeness (dark green is 90+%). The circles are the last week of seismicity from the NEIC significant earthquakes json feed.

Data Access Tools Accessible to NEIC

The NEIC has direct access to the server through an alternate port SSH. Within the NEIC's user directory, the primary data access scripts are linked directly to the master scripts written by the PI. The PI has also added these scripts to a GitHub repository (https://github.com/crowellbw/neic). The primary scripts are ***get_timeseries*** and ***get_pgd***, both written in Python3. Both of these scripts connect to MongoDB to obtain data based on user inputs. There are three primary modes of data access:

1) Using an existing USGS eventid - This will call the Python function get_usgsevent.eventinformation(eventid) and retrieve the origin time, the hypocenter and the magnitude from the M>4.5 weekly geojson feed from the USGS. Once an event is found, the user will be asked for a search radius for stations and the amount of data, in seconds, you would like to retrieve. A final prompt asks if the user would like to create GMT plots of the time series (***get_timeseries***) or the peak ground displacements (***get_pgd***). This will also retrieve 60 seconds of pre-event signal. Text files and .eps plots are created in the user's home directory under ~/gnssoutput/eventid/.

2) Using manually entered event information - This mode prompts the user to enter in the origin information (UTC origin time and hypocenter) before prompting for search radius and amount of data to access. The user needs to assign an event id for this mode, but it operates exactly the same as using an eventid. The final files are stored using the same directory structure (~/gnssoutput/eventid/).
3) A single station between two epochs of time - This mode only exists for get_timeseries. The prompt is more simple, just a 4 character station id, the start time, and the number of seconds to download. Rather than specifying an eventid, the time series file and plots are stored in a folder named after the day within the gnssoutput folder.

Example Workflow for Rapid Analysis

Since this service has been operational, only two earthquakes large enough to be recorded on real-time GNSS occurred, the Mw 6.5 Puerto Rico earthquake on January 7, 2020, and the Mw 7.7 Lucea, Jamaica earthquake on January 28, 2020. Unfortunately, few stations were in the near-field of these events, so we did not obtain meaningful PGD values or coseismic offsets. However, for the Lucea earthquake, we ran the ***get_timeseries*** script with a search radius of 1000 km and resolved a directivity pulse at station UNPM in Cancun, Mexico, 860 km NW of the source. Figure 2 shows the plot generated by ***get_timeseries*** showing a strong Love wave arrival at the station in the north component.

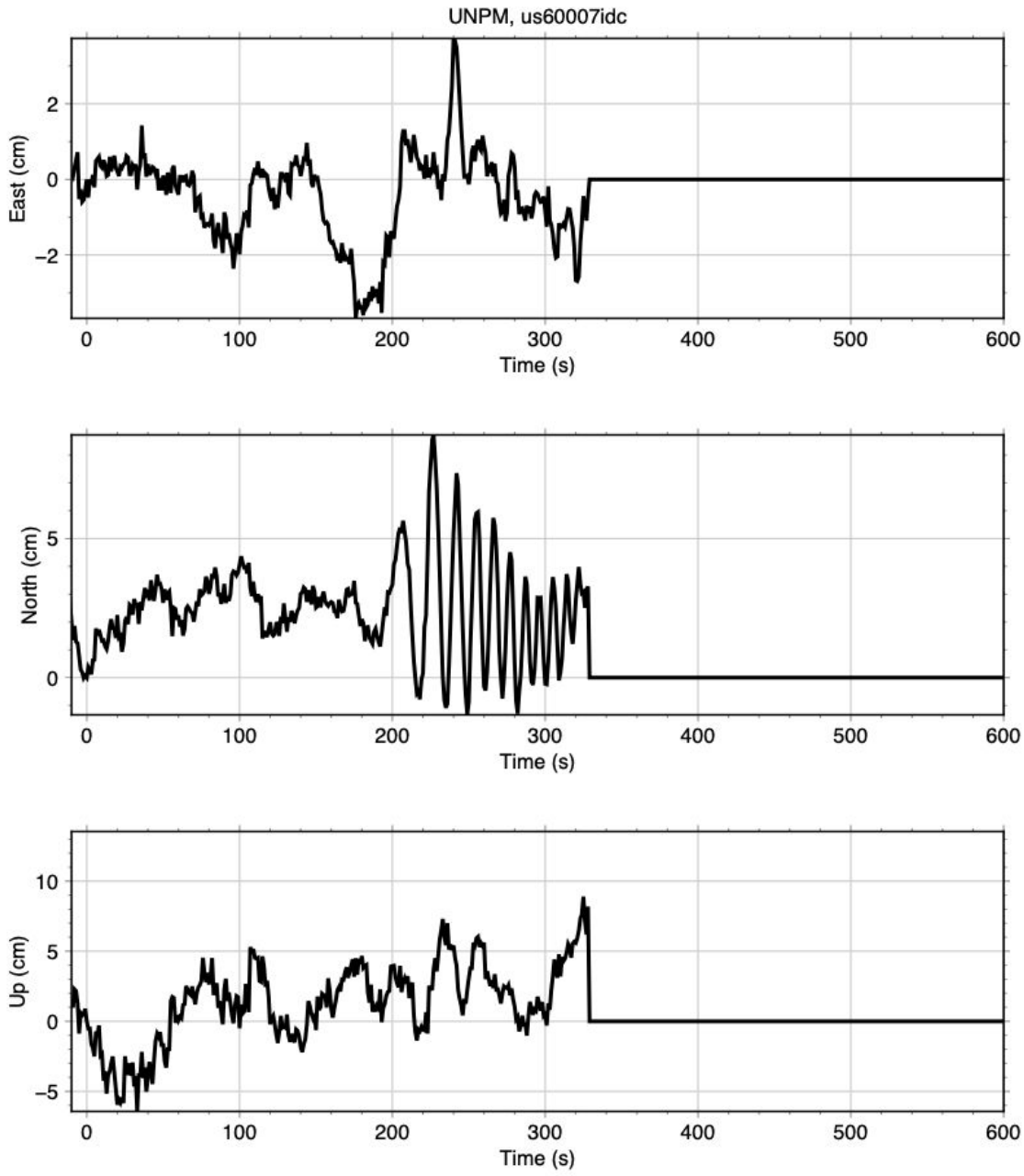

Figure 2: Real-time GNSS data recorded during the January 28, 2020 Mw 7.7 Lucea, Jamaica earthquake in Cancun, Mexico, 860 km NW of the source. In the north component (center), a strong Love wave arrival can be observed, undulating for over 100 seconds. At 320 seconds, we lose signal at the station due to either power or telemetry issues at the station.

On March 18, 2020, the Mw 5.7 Magna, Utah earthquake struck just outside Salt Lake City. While we do not expect to see any signals from this event, we ran the ***get_pgd*** script to retrieve all the stations within 500 km. Figure 3 shows the PGD scaling result for 32 stations. Most of the stations are under 10 cm and well under the critical PGD value theoretically derived

by Melgar et al. [2019]. There are some outliers for further stations that we need to address, however, if an analyst ran the ***get_timeseries*** script and looked at individual waveforms, they can determine that no appreciable signal is observed.

The way in which the PI has been using this service to date is as follows:

1) Once an alert of an earthquake is received, he checks the leaflet map to see how many stations are nearby.
2) If there is a station within a few hundred kilometers, he will run ***get_timeseries*** to query the database to see if data exists during the event and to look at the plots.
3) If there are signals visible, he runs ***get_pgd*** to obtain a pgd overview file and a plot of the observations.
4) If there were signals available, he informs the NEIC and other interested parties.
5) If there are offsets, he runs these through the offline version of G-FAST to obtain a slip model and CMT solution
6) He begins a post-processing run with either TRACK or GipsyX to further refine the solutions.

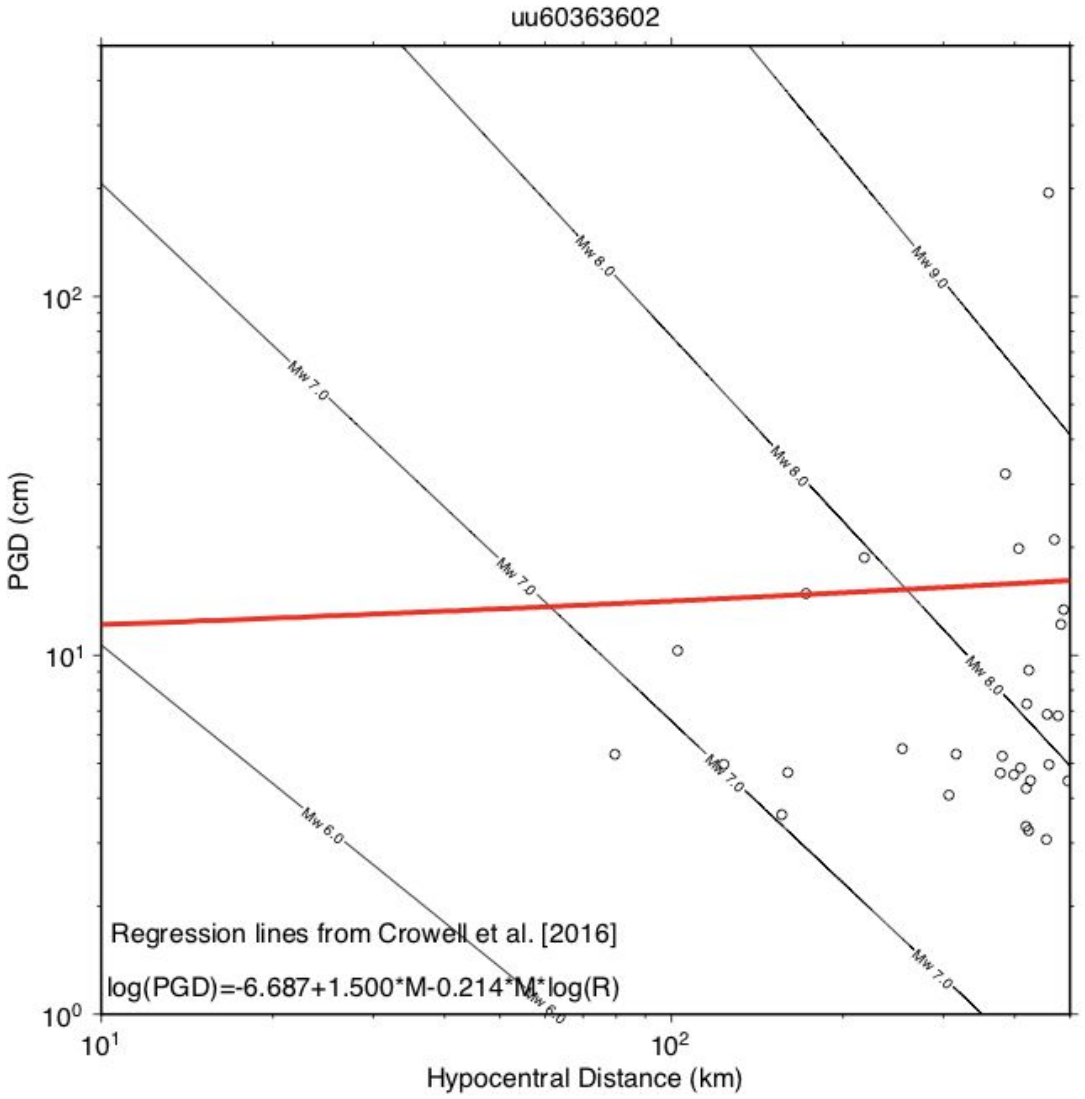

Figure 3: PGD results for the March 18, 2020 Mw 5.7 Magna, Utah earthquake. The red line is the critical PGD value from Melgar et al. [2019], indicating that values below this line are within the noise. We would not expect this earthquake to produce noticeable signals on the GNSS time series as it is too small, so this represents only noise.

Next Steps

The NEIC has full access to the UW database server, geodesy.ess.washington.edu, and has read/write access to MongoDB. Only base level tools were built for this project for data access and manipulation, and it will be the job of an incoming Mendenhall postdoctoral scholar to link the data directly into the NEIC's Hydra system. PI Crowell will continue to maintain the server and build additional tools for data access and visualization. A key next step will be to

build web services that use the data access tools for a much easier way to access and manipulate the data.

As the web services are further refined, we expect to share this service with other seismic analysis centers around the world. We are currently working with GNS in New Zealand and EPN in Ecuador to provide these solutions back to them. As we advocate and present this system to other groups, we expect data access will be opened up, further improving the system for all. This same data stream is also being broadcast to NOAA's Tsunami Warning Centers for a separate NASA funded project to add GNSS source products into the warning chain.

Project Data

All data produced in this project is publically available. The web service portal at http://geodesy.ess.washington.edu/data currently hosts the processed position streams. Level 0 raw GNSS data in Rinex format is available from several data providers such as UNAVCO, SOPAC, PANGA, etc. The RabbitMQ streams are accessible by request to CWU.

Bibliography

No publications have been made as a result of this project as of the writing of this technical final report. A publication is being prepared by the PI that will be submitted in late 2020.

References

Blewitt, G., C. Kreemer, W. C. Hammond, H.‐P. Plag, S. Stein, and E. Okal (2006), Rapid determination of earthquake magnitude using GPS for tsunami warning systems, *Geophys. Res. Lett.*, 33, L11309, doi:10.1029/2006GL026145.

Crowell, B.W., Y. Bock, and M.B. Squibb (2009), Demonstration of earthquake early warning using total displacement waveforms from real-time GPS networks, *Seism. Res. Lett.*, 80 (5), 768-778, doi:10.1785/gssrl.80.5.772.

Crowell, B.W., D. Melgar, Y. Bock, J.S. Haase, and J. Geng (2013), Earthquake magnitude scaling using seismogeodetic data, *Geophys. Res. Lett.*, 40, 6089-6094, doi: 10.1002/2013GL058391.

Crowell, B. W., D. A. Schmidt, P. Bodin, J. E. Vidale, B. Baker, S. Barrientos, and J. Geng (2018a), G-FAST earthquake early warning potential for great earthquakes in Chile, *Seism. Res. Lett.*, doi: 10.1785/0220170180.

Crowell, B. W., D. Melgar, and J. Geng (2018b), Hypothetical real-time GNSS modeling of the 2016 Mw 7.8 Kaikoura Earthquake: Perspectives from ground motion and tsunami inundation prediction, *Bull. Seism. Soc. Am.*, doi:10.1785/0120170247.

Duputel, Z., L. Rivera, H. Kanamori, G. P. Hayes, B. Hirshorn, and S. Weinstein (2011), Real-time W phase inversion during the 2011 off the Pacific coast of Tohoku earthquake, *Earth Planets Space*, 63, 535-539, doi: 10.5047/eps.2011.05.032.

Geng, J., Y. Bock, D. Melgar, B.W. Crowell, and J.S. Haase (2013), A new seismogeodetic approach applied to GPS and accelerometer observations of the 2012 Brawley seismic swarm: Implications for earthquake early warning, *Geochem. Geophys. Geosyst.*, 14, doi: 10.1002/ggge.20144.

Genrich, J. F., and Y. Bock (2006), Instantaneous geodetic positioning with 10–50 Hz GPS measurements: Noise characteristics and implications for monitoring networks, *J. Geophys. Res.*, 111, B03403, doi: 10.1029/2005JB003617.

Grapenthin, R., M. West, and J. Freymueller (2017), The utility of GNSS for earthquake early warning in regions with sparse seismic networks, *Bull. Seism. Soc. Am.,* 107 (4), 1883–1890, doi: 10.1785/0120160317.

Hayes, G. P., et al. (2011), 88 Hours: The U.S. Geological Survey National Earthquake Information Center response to the 11 March 2011 Mw 9.0 Tohoku earthquake, *Seism. Res. Lett*., 82, 481-493, doi: 10.1785/gssrl.82.4.481.

Hoshiba, M., et al. (2011), Outline of the 2011 off the Pacific coast of Tohoku earthquake (Mw 9.0) - Earthquake early warning and observed seismic intensity, *Earth Planets Space*, 63, 547-551, doi: 10.5047/eps.2011.05.031.

Kouba, J., and P. Héroux (2001), Precise Point Positioning Using IGS Orbit and Clock Products *GPS Solutions*, 5:12, doi:10.1007/PL00012883.

Kanamori, H., and L. Rivera (2008), Source inversion of W phase: Speeding up seismic tsunami warning. Geophys. J. Int., 175, 222–238.

Melgar, D., Y. Bock, D. Sanchez, and B.W. Crowell (2013), On robust and reliable automated baseline corrections for strong motion seismology, *J. Geophys. Res.*, 118, doi: 10.1002/jgrb.50135.

Melgar, D., J. Geng, B. W. Crowell, J. S. Haase, Y. Bock, W. C. Hammond, and R. M. Allen (2015a), Seismogeodesy of the 2014 Mw6.1 Napa Earthquake, California: Rapid Response and Modeling of Fast Rupture on a Dipping Strike-slip fault, *J. Geophys. Res.*, 120, doi: 10.1002/2015JB011921.

Melgar, D., B. W. Crowell, J. Geng, R. M. Allen, Y. Bock, S. Riquelme, E. M. Hill, M. Protti, and A. Ganas (2015b), Earthquake magnitude calculation without saturation from the scaling of peak ground displacement, *Geophys. Res. Lett.*, 42, 5197-5205, doi: 10.1002/2015GL064278.

Melgar, D., et al. (2016), Local tsunami warnings: Perspectives from recent large events, *Geophys. Res. Lett.*, 43, 1109–1117, doi: 10.1002/2015GL067100.

Melgar, D., T. I. Melbourne, B. W. Crowell, J. Geng, W. Szeliga, C. Scrivner, M. Santillan, and D. E. Goldberg (2019), Real-time high-rate GNSS displacements: Performance demonstration during the 2019 Ridgecrest, CA earthquakes, *Seism. Res. Lett.,* doi: 10.1785/0220190223.

Polet, J., and H. K. Thio (2011), Rapid calculation of a Centroid Moment Tensor and waveheight predictions around the north Pacific for the 2011 off the Pacific coast of Tohoku earthquake, *Earth Planets Space*, 63, 541-545, doi:10.5047/eps.2011.05.005.

Santillan, V. M., Melbourne, T. I., Szeliga, W. M., Scrivner, C. W. (2013), A fast-convergence stream editor for real-time precise point positioning, Annual Meeting of the American Geophysical Union, Paper No. G53B–0930, San Francisco, CA.